

1652

RAW SEQUENCE LISTING                      DATE: 06/14/2000  
PATENT APPLICATION:    US/09/402,488A              TIME: 15:32:14

Input Set : A:\Sequence  
Output Set: N:\CRF3\06142000\I402488A.raw

**ENTERED****SEQUENCE LISTING**

3 (1) GENERAL INFORMATION:  
C--> 5 (i) APPLICANT: van Rooijen, Gijs  
6 Alcantara, Joenel  
7 Moloney, Maurice M.  
9 (ii) TITLE OF INVENTION: Method for Cleavage of Fusion Proteins  
11 (iii) NUMBER OF SEQUENCES: 4  
13 (iv) CORRESPONDENCE ADDRESS:  
14 (A) ADDRESSEE: BERESKIN & PARR  
15 (B) STREET: 40 King Street West  
16 (C) CITY: Toronto  
17 (D) STATE: Ontario  
18 (E) COUNTRY: Canada  
19 (F) ZIP: M5H 3Y2  
21 (v) COMPUTER READABLE FORM:  
22 (A) MEDIUM TYPE: Floppy disk  
23 (B) COMPUTER: IBM PC compatible  
24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30  
C--> 27 (vi) CURRENT APPLICATION DATA:  
C--> 28 (A) APPLICATION NUMBER: US/09/402,488A  
C--> 29 (B) FILING DATE: 16-Feb-2000  
30 (C) CLASSIFICATION:  
32 (viii) ATTORNEY/AGENT INFORMATION:  
33 (A) NAME: Gravelle, Micheline  
34 (B) REGISTRATION NUMBER: 40,261  
35 (C) REFERENCE/DOCKET NUMBER: 9369-98  
37 (ix) TELECOMMUNICATION INFORMATION:  
38 (A) TELEPHONE: (416) 364-7311  
39 (B) TELEFAX: (416) 361-1398  
42 (2) INFORMATION FOR SEQ ID NO: 1:  
44 (i) SEQUENCE CHARACTERISTICS:  
45 (A) LENGTH: 1096 base pairs  
46 (B) TYPE: nucleic acid  
47 (C) STRANDEDNESS: single  
48 (D) TOPOLOGY: linear  
50 (ii) MOLECULE TYPE: other nucleic acid  
53 (ix) FEATURE:  
54 (A) NAME/KEY: CDS  
55 (B) LOCATION: 1..1029  
58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
60 ATG TCC CCT ATA CTA GGT TAT TGG AAA ATT AAG GGC CTT GTG CAA CCC 48  
61 Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro  
62 1 5 10 15  
64 ACT CGA CTT CTT TTG GAA TAT CTT GAA GAA AAA TAT GAA GAG CAT TTG 96  
65 Thr Arg Leu Leu Leu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu  
66 20 25 30

RAW SEQUENCE LISTING                      DATE: 06/14/2000  
 PATENT APPLICATION:    US/09/402,488A            TIME: 15:32:14

Input Set : A:\Sequence  
 Output Set: N:\CRF3\06142000\I402488A.raw

68	TAT GAG CGC GAT GAA GGT GAT AAA TGG CGA AAC AAA AAG TTT GAA TTG	144
69	Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu	
70	35 40 45	
72	GGT TTG GAG TTT CCC AAT CTT CCT TAT TAT ATT GAT GGT GAT GTT AAA	192
73	Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys	
74	50 55 60	
76	TTA ACA CAG TCT ATG GCC ATC ATA CGT TAT ATA GCT GAC AAG CAC AAC	240
77	Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn	
78	65 70 75 80	
80	ATG TTG GGT GGT TGT CCA AAA GAG CGT GCA GAG ATT TCA ATG CTT GAA	288
81	Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu	
82	85 90 95	
84	GGA GCG GTT TTG GAT ATT AGA TAC GGT GTT TCG AGA ATT GCA TAT AGT	336
85	Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser	
86	100 105 110	
88	AAA GAC TTT GAA ACT CTC AAA GTT GAT TTT CTT AGC AAG CTA CCT GAA	384
89	Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu	
90	115 120 125	
92	ATG CTG AAA ATG TTC GAA GAT CGT TTA TGT CAT AAA ACA TAT TTA AAT	432
93	Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn	
94	130 135 140	
96	GGT GAT CAT GTA ACC CAT CCT GAC TTC ATG TTG TAT GAC GCT CTT GAT	480
97	Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp	
98	145 150 155 160	
100	GTT GTT TTA TAC ATG GAC CCA ATG TGC CTG GAT GCG TTC CCA AAA TTA	528
101	Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu	
102	165 170 175	
104	GTT TGT TTT AAA AAA CGT ATT GAA GCT ATC CCA CAA ATT GAT AAG TAC	576
105	Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr	
106	180 185 190	
108	TTG AAA TCC AGC AAG TAT ATA GCA TGG CCT TTG CAG GGC TGG CAA GCC	624
109	Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala	
110	195 200 205	
112	ACG TTT GGT GGT GGC GAC CAT CCT CCA AAA TCG GAT CTG GTT CCG CGT	672
113	Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg	
114	210 215 220	
116	GGA TCC CCG AAT TCC CGG GTC GAC TCG AGC GGC GCT GAG ATC ACC	720
117	Gly Ser Pro Asn Ser Arg Val Asp Ser Ser Gly Arg Ala Glu Ile Thr	
118	225 230 235 240	
120	AGG ATC CCT CTG TAC AAA GGC AAG TCT CTG AGG AAG GCG CTG AAG GAG	768
121	Arg Ile Pro Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu	
122	245 250 255	
124	CAT GGG CTT CTG GAG GAC TTC CTG CAG AAA CAG CAG TAT GGC ATC AGC	816
125	His Gly Leu Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser	
126	260 265 270	
128	AGC AAG TAC TCC GGC TTC GTC GTC TAT ACC GAC TGT ACC GAG TCC GGT	864
129	Ser Lys Tyr Ser Gly Phe Val Val Tyr Thr Asp Cys Thr Glu Ser Gly	
130	275 280 285	
132	CAG AAC CTC TGT CTC TGT GAG GGT TCC AAC GTC TGT GGT CAG GGT AAC	912

## RAW SEQUENCE LISTING

DATE: 06/14/2000

PATENT APPLICATION: US/09/402,488A

TIME: 15:32:14

Input Set : A:\Sequence

Output Set: N:\CRF3\06142000\I402488A.raw

```

133 Gln Asn Leu Cys Leu Cys Glu Gly Ser Asn Val Cys Gly Gln Gly Asn
134      290      295      300
136 AAG TGT ATC CTC GGT TCC GAC GGT GAG AAG AAC CAG TGT GTC ACC GGT      960
137 Lys Cys Ile Leu Gly Ser Asp Gly Glu Lys Asn Gln Cys Val Thr Gly
138 305      310      315      320
140 GAG GGA ACC CCA AAG CCA CAG TCC CAC AAC GAC GGT GAC TTT GAG GAG      1008
141 Glu Gly Thr Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe Glu Glu
142      325      330      335
144 ATC CCA GAG GAG TAT CTC CAG TAAAGATCTA AGCTTGCTGC TGCTATCGAA      1059
145 Ile Pro Glu Glu Tyr Leu Gln
146      340
148 TTCCTGCAGC CCGGGGGATC CACTAGTTCT AGAGCGG      1096
151 (2) INFORMATION FOR SEQ ID NO: 2:
153 (i) SEQUENCE CHARACTERISTICS:
154 (A) LENGTH: 343 amino acids
155 (B) TYPE: amino acid
156 (D) TOPOLOGY: linear
158 (ii) MOLECULE TYPE: protein
160 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
162 Met Ser Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val Gln Pro
163 1      5      10      15
165 Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu His Leu
166      20      25      30
168 Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe Glu Leu
169      35      40      45
171 Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp Val Lys
172      50      55      60
174 Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys His Asn
175 65      70      75      80
177 Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met Leu Glu
178      85      90      95
180 Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala Tyr Ser
181      100      105      110
183 Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu Pro Glu
184      115      120      125
186 Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr Leu Asn
187      130      135      140
189 Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala Leu Asp
190 145      150      155      160
192 Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro Lys Leu
193      165      170      175
195 Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp Lys Tyr
196      180      185      190
198 Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp Gln Ala
199      195      200      205
201 Thr Phe Gly Gly Gly Asp His Pro Pro Lys Ser Asp Leu Val Pro Arg
202      210      215      220
204 Gly Ser Pro Asn Ser Arg Val Asp Ser Ser Gly Arg Ala Glu Ile Thr
205 225      230      235      240

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/402,488A DATE: 06/14/2000  
 TIME: 15:32:14

Input Set : A:\Sequence  
 Output Set: N:\CRF3\06142000\I402488A.raw

```

207 Arg Ile Pro Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu
208                               245 250 255
210 His Gly Leu Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser
211                               260 265 270
213 Ser Lys Tyr Ser Gly Phe Val Val Tyr Thr Asp Cys Thr Glu Ser Gly
214                               275 280 285
216 Gln Asn Leu Cys Leu Cys Glu Gly Ser Asn Val Cys Gly Gln Gly Asn
217                               290 295 300
219 Lys Cys Ile Leu Gly Ser Asp Gly Glu Lys Asn Gln Cys Val Thr Gly
220 305                               310 315 320
222 Glu Gly Thr Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe Glu Glu
223                               325 330 335
225 Ile Pro Glu Glu Tyr Leu Gln
226                               340
228 (2) INFORMATION FOR SEQ ID NO: 3:
230 (i) SEQUENCE CHARACTERISTICS:
231 (A) LENGTH: 819 base pairs
232 (B) TYPE: nucleic acid
233 (C) STRANDEDNESS: single
234 (D) TOPOLOGY: linear
236 (ii) MOLECULE TYPE: other nucleic acid
239 (ix) FEATURE:
240 (A) NAME/KEY: CDS
241 (B) LOCATION: 1..816
244 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
246 ATG CGG GGT TCT CAT CAT CAT CAT CAT GGT ATG GCT AGC ATG ACT 48
247 Met Arg Gly Ser His His His His His Gly Met Ala Ser Met Thr
248 1 5 10 15
250 GGT GGA CAG CAA ATG GGT CGG GAT CTG TAC GAC GAT GAC GAT AAG GAT 96
251 Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Lys Asp
252 20 25 30
254 CCG AGC TCG AGA TCT GCA GAA ATC GGA TCC GCT GAG ATC ACC AGG ATC 144
255 Pro Ser Ser Arg Ser Ala Glu Ile Gly Ser Ala Glu Ile Thr Arg Ile
256 35 40 45
258 CCT CTG TAC AAA GGC AAG TCT CTG AGG AAG GCG CTG AAG GAG CAT GGG 192
259 Pro Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly
260 50 55 60
262 CTT CTG GAG GAC TTC CTG CAG AAA CAG CAG TAT GGC ATC AGC AGC AAG 240
263 Leu Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys
264 65 70 75 80
266 TAC TCC GGC TTC TCA GAC AAC CAG CGG CTC TTC AAT AAT GCA GTC ATT 288
267 Tyr Ser Gly Phe Ser Asp Asn Gln Arg Leu Phe Asn Asn Ala Val Ile
268 85 90 95
270 CGT GTA CAA CAC CTG CAC CAG CTG GCT GCA AAA ATG ATT AAC GAC TTT 336
271 Arg Val Gln His Leu His Gln Leu Ala Ala Lys Met Ile Asn Asp Phe
272 100 105 110
274 GAG GAC AGC CTG TTG CCT GAG GAA CGC AGA CAG CTG AGT AAA ATC TTC 384
275 Glu Asp Ser Leu Leu Pro Glu Glu Arg Arg Gln Leu Ser Lys Ile Phe
276 115 120 125

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/402,488A DATE: 06/14/2000  
 TIME: 15:32:14

Input Set : A:\Sequence  
 Output Set: N:\CRF3\06142000\I402488A.raw

```

278 CCT CTG TCT TTC TGC AAT TCT GAC TAC ATT GAG GCG CCT GCT GGA AAA      432
279 Pro Leu Ser Phe Cys Asn Ser Asp Tyr Ile Glu Ala Pro Ala Gly Lys
280      130      135      140
282 GAT GAA ACA CAG AAG AGC TCT ATG CTG AAG CTT CTT CGC ATC TCT TTT      480
283 Asp Glu Thr Gln Lys Ser Ser Met Leu Lys Leu Arg Ile Ser Phe
284 145      150      155      160
286 CAC CTC ATT GAG TCC TGG GAG TTC CCA AGC CAG TCC CTG AGC GGA ACC      528
287 His Leu Ile Glu Ser Trp Glu Phe Pro Ser Gln Ser Leu Ser Gly Thr
288      165      170      175
290 GTC TCA AAC AGC CTG ACC GTA GGG AAC CCC AAC CAG CTC ACT GAG AAG      576
291 Val Ser Asn Ser Leu Thr Val Gly Asn Pro Asn Gln Leu Thr Glu Lys
292      180      185      190
294 CTG GCC GAC TTG AAA ATG GGC ATC AGT GTG CTC ATC CAG GCA TGT CTC      624
295 Leu Ala Asp Leu Lys Met Gly Ile Ser Val Leu Ile Gln Ala Cys Leu
296      195      200      205
298 GAT GGT CAA CCA AAC ATG GAT GAT AAC GAC TCC TTG CCG CTG CCT TTT      672
299 Asp Gly Gln Pro Asn Met Asp Asp Asn Asp Ser Leu Pro Leu Pro Phe
300      210      215      220
302 GAG GAC TTC TAC TTG ACC ATG GGG GAG AAC AAC CTC AGA GAG AGC TTT      720
303 Glu Asp Phe Tyr Leu Thr Met Gly Glu Asn Asn Leu Arg Glu Ser Phe
304 225      230      235      240
306 CGT CTG CTG GCT TGC TTC AAG AAG GAC ATG CAC AAA GTC GAG ACC TAC      768
307 Arg Leu Leu Ala Cys Phe Lys Lys Asp Met His Lys Val Glu Thr Tyr
308      245      250      255
310 TTG AGG GTT GCA AAT TGC AGG AGA TCC CTG GAT TCC AAC TGC ACC CTG      816
311 Leu Arg Val Ala Asn Cys Arg Arg Ser Leu Asp Ser Asn Cys Thr Leu
312      260      265      270
314 TAG      819
317 (2) INFORMATION FOR SEQ ID NO: 4:
319 (i) SEQUENCE CHARACTERISTICS:
320 (A) LENGTH: 272 amino acids
321 (B) TYPE: amino acid
322 (D) TOPOLOGY: linear
324 (ii) MOLECULE TYPE: protein
326 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
328 Met Arg Gly Ser His His His His His Gly Met Ala Ser Met Thr
329 1      5      10      15
331 Gly Gly Gln Gln Met Gly Arg Asp Leu Tyr Asp Asp Asp Asp Lys Asp
332      20      25      30
334 Pro Ser Ser Arg Ser Ala Glu Ile Gly Ser Ala Glu Ile Thr Arg Ile
335      35      40      45
337 Pro Leu Tyr Lys Gly Lys Ser Leu Arg Lys Ala Leu Lys Glu His Gly
338      50      55      60
340 Leu Leu Glu Asp Phe Leu Gln Lys Gln Gln Tyr Gly Ile Ser Ser Lys
341 65      70      75      80
343 Tyr Ser Gly Phe Ser Asp Asn Gln Arg Leu Phe Asn Asn Ala Val Ile
344      85      90      95
346 Arg Val Gln His Leu His Gln Leu Ala Ala Lys Met Ile Asn Asp Phe
347      100      105      110

```

VERIFICATION SUMMARY                      DATE: 06/14/2000  
PATENT APPLICATION:    US/09/402,488A       TIME: 15:32:15

Input Set : A:\Sequence  
Output Set: N:\CRF3\06142000\I402488A.raw

L:5 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]